IN THE CLAIMS:

The following is a complete listing of the claims, and replaces all earlier listings and all earlier versions.

1.   (Currently Amendment) A document segmentation apparatus for segmenting a list type table for describing data or a layout type table represented between tags defined in a language for use in composing web pages, comprising:

table analyzing means for generating cell vectors representing characteristics of cells having at least one of row width and column width of the cells, and cell position data indicating a positional relationship between cells and cell vectors representing characteristics of the cells, by analyzing a table pinched between a start tag and an end tag in a document to be processed;

table type judging means for judging a table type whether the table analyzed by said table analyzing means is a list type table for describing data or a layout type table for describing a layout of page with reference to the cell position data and the cell vectors generated by said table analyzing means;

first segment generating means for generating a segment plurality of segments each of which is pinched between that start tag and the end tag by dividing the table with a first method in a case in which said table type judging means judges that the table [[type]] analyzed by said table analyzing means is [[a]] the list type table; and

second segment generating means for generating a plurality of segments each of which is pinched between the start tag and the end tag by dividing the table with a second

- 2 -

method in a case in which <u>said table type judging means judges that</u> the table [[type]] <u>analyzed</u> <u>by said table analyzing means</u> is [[a]] <u>the</u> layout type <u>table</u>.

2.   (Previously Presented) A document segmentation apparatus according to claim 1, wherein said first segment generating means comprise:

cut direction determination means for determining a cut direction of the table by judging whether the data is expressed in a column or a row in the table on the basis of the cell position data and the cell vectors; and

table segment generating means for generating a table segment by dividing the table on the basis of the table type and the cut direction.

3.   (Original) A document segmentation apparatus according to claim 2, wherein said second segment generating means generate the table itself as the segment.

4.   (Previously Presented) A document segmentation apparatus according to claim 1, wherein said second segment generating means comprise:

cell cluster generating means for generating cell cluster information by clustering the cells in the table; and

layout segment generating means for generating segment by connecting the cells in the table with reference to the cell position data and the cell cluster information.

- 3 -

5.   (Original) A document segmentation apparatus according to claim 4, wherein said first segment generating means generate the table itself as the segment.

6.   (Original) A document segmentation apparatus according to claim 4, wherein said second segment generating means generate the table itself as the segment.

7.   (Previously Presented) A document segmentation apparatus according to claim 1, further comprising normal segment generating means for dividing the document into a segment which corresponds to one table,

wherein the table generated as one segment by said normal segment generating means is to be processed by said table analyzing means.

8.   (Original) A document segmentation apparatus according to claim 1, wherein said table analyzing means further generate cell data of the analyzed table and said table and said table type judging means judge the table type with reference to the cell data.

9.   (Original) A document segmentation apparatus according to claim 8, wherein said table type judging means comprise similarity judging means for judging the table type on the basis of similarity between the cell data positioned at particular positions with reference to the cell position data and the cell data generated by said table analyzing means.

- 4 -

10.   (Original) A document segmentation apparatus according to claim 8, wherein said table type judging means comprise partial character line extracting means for extracting partial character lines from the cell data positioned at a particular position with reference to the cell position data and the cell data generated by said table analyzing means, and character line comparing means for comparing the extracted partial character lines to judge the table type.

11.   (Original) A document segmentation apparatus according to claim 8, wherein said table type judging means comprise partial character line extracting means for extracting partial character lines from the cell data positioned at a particular position with reference to the cell position data and the cell data generated by said table analyzing means, and similarity judging means for judging the table type on the basis of similarity between the extracted partial character lines.

12.   (Original) A document segmentation apparatus according to claim 8, wherein said table type judging means comprise syntax judging means for judging the table type with reference to the cell position data, the cell vectors and the cell data generated by said table analyzing means, and similarity judging means for judging the table type on the basis of similarity between the cell data positioned at particular positions with reference to the cell position data and the cell data generated by said table analyzing means.

- 5 -

13.   (Original) A document segmentation apparatus according to claim 8, wherein said table type judging means comprise syntax judging means for judging the table type with reference to the cell position data, the cell vectors and the cell data generated by said table analyzing means, partial character line extracting means for extracting partial character lines from the cell data positioned at a particular position with reference to the cell position data and the cell data generated by said table analyzing means, and, character line comparing means for comparing the extracted partial character lines to judge the table type.

14.   (Original) A document segmentation apparatus according to claim 8, wherein said table type judging means comprise syntax judging means for judging the table type with reference to the cell position data, the cell vectors and the cell data generated by said table analyzing means, partial character line extracting means for extracting partial character lines from the cell data positioned at a particular position with reference to the cell position data and the cell data generated by said table analyzing means, and similarity judging means for judging the table type on the basis of similarity between the extracted partial character lines.

15.   (Previously Presented) A document segmentation apparatus according to claim 1, further comprising table reforming means for reforming the table so that the number of cells in each column and each row becomes the same, by analyzing the table to be processed,

wherein said table analyzing means analyze the reformed table.

- 6 -

16.   (Original) A document segmentation apparatus according to claim 15, wherein said table reforming means comprise supplementary data removing means for removing data added to the table from the table data.

17.   (Original) A document segmentation apparatus according to claim 15, wherein said table reforming means comprise multi-row/multi-column processing means for reforming the table regularly by analyzing the structure of the table data.

18.   (Original) A document segmentation apparatus according to claim 15, wherein said table reforming means comprise composite table processing means for reforming the table by analyzing regularity of information description constituting the table.

19.   (Previously Presented) A document segmentation apparatus according to claim 15, wherein said table reforming means comprise:

supplementary data removing means for removing data added to the table from the table data; and

multi-row/multi-column processing means for reforming the table regularly by analyzing the structure of the table data.

20.   (Previously Presented) A document segmentation apparatus according to claim 15, wherein said table reforming means comprise:

- 7 -

supplementary data removing means for removing data added to the table from the table data; and

composite table processing means for reforming the table by analyzing regularity of information description constituting the table.

21.   (Previously Presented) A document segmentation apparatus according to claim 15, wherein said table reforming means comprise:

multi-row/multi-column processing means for reforming the table regularly by analyzing the structure of the table data; and

composite table processing means for reforming the table by analyzing regularity of information description constituting the table.

22.   (Original) A document segmentation apparatus according to claim 15, wherein said table reforming means comprise:

supplementary data removing means for removing data added to the table from the table data;

multi-row/multi-column processing means for reforming the table regularly by analyzing the structure of the table data; and

composite table processing means for reforming the table by analyzing regularity of information description constituting the table.

- 8 -

23.   (Currently Amended) A document segmentation method for segmenting

a list type table for describing data or a layout type table representing between tags defined in

a language for use in composing web pages, comprising:

a table analyzing step, of generating cell vectors representing characteristics of

cells having at least one of row width and column width of the cells, and cell position data

indicating a positional relationship between cells and cell vectors representing characteristics

of the cells, by analyzing a table pinched between a start tag and an end tag in a document to

be processed;

a table type judging step, of judging a table type whether the table analyzed in

said table analyzing step is a list type table for describing data or a layout type table for

describing a layout of a page with reference to the cell position data and the cell vectors

generated in said table analyzing step;

a first segment generating step, of generating a plurality of segments each of

which is pinched between the start tag and the end tag by dividing the table with a first method

in a case in which it is judged that the table [[type]] analyzed in said table analyzing step is a

list format type table in said table type judging step; and

a second segment generating step, of generating a plurality of segments each

of which is pinched between the start tag and the end tag by dividing the table with a second

method in a case in which it is judged that the table type analyzed in said table analyzing step

is a layout type table in said table type judging step.

-9-

24.   (Previously Presented) A document segmentation method according to claim 23, wherein said first segment generating step comprises:

a cut direction determination step, of determining a cut direction of the table by judging whether the data is expressed in a column or a row in the table on the basis of the cell position data and the cell vectors; and

a table segment generating step, of generating a table segment by dividing the table on the basis of the table type and the cut direction.

25.   (Previously Presented) A document segmentation method according to claim 24, wherein said second segment generating step includes generating the table itself as the segment.

26.   (Previously Presented) A document segmentation method according to claim 23, wherein said second segment generating step comprises:

a cell cluster generating step, of generating cell cluster information by clustering the cells in the table; and

a layout segment generating step, of generating segment by connecting the cells in the table with reference to the cell position data and the cell cluster information.

27.   (Previously Presented) A document segmentation method according to claim 26, wherein said first segment generating step includes generating the table itself as the segment.

- 10 -

28.   (Previously Presented) A document segmentation method according to claim 26, wherein said second segment generating step includes generating the table itself as the segment.

29.   (Previously Presented) A document segmentation method according to claim 23, further comprising a normal segment generating step, of dividing the document into a segment which corresponds to one table,

        wherein the table generated as one segment in said normal segment generating step is to be processed in said table analyzing step.

30.   (Previously Presented) A document segmentation method according to claim 23, wherein said table analyzing step further includes generating cell data of the analyzed table and said table type judging step includes judging the table type with reference to the cell data.

31.   (Previously Presented) A document segmentation method according to claim 30, wherein said table type judging step comprises a similarity judging step, of judging the table type on the basis of similarity between the cell data positioned at particular positions with reference to the cell position data and the cell data generated in said table analyzing step.

32.   (Previously Presented) A document segmentation method according to claim 30, wherein said table type judging step comprises a partial character line extracting

- 11 -

step, of extracting partial character lines from the cell data positioned at a particular position

with reference to the cell position data and the cell data generated in said table analyzing step,

and a character line comparing step, of comparing the extracted partial character lines to judge

the table type.

33.   (Previously Presented) A document segmentation method according to

claim 30, wherein said table type judging step comprises a partial character line extracting

step, of extracting partial character lines from the cell data positioned at a particular position

with reference to the cell position data and the cell data generated in said table analyzing step,

and a similarity judging step, of judging the table type on the basis of similarity between the

extracted partial character lines.

34.   (Previously Presented) A document segmentation method according to

claim 30, wherein said table type judging step comprises a syntax judging step, of judging the

table type with reference to the cell position data, the cell vectors and the cell data generated

in said table analyzing step, and a similarity judging step, of judging the table type on the basis

of similarity between the cell data positioned at particular positions with reference to the cell

position data and the cell data generated in said table analyzing step.

35.   (Previously Presented) A document segmentation method according to

claim 30, wherein said table type judging step comprises a syntax judging step, of judging the

table type with reference to the cell position data, the cell vectors and the cell data generated

- 12 -

in said table analyzing step, a partial character line extracting step, of extracting partial character lines from the cell data positioned at a particular position with reference to the cell position data and the cell data generated in said table analyzing step, and a character line comparing step, of comparing the extracted partial character lines to judge the table type.

36.   (Previously Presented) A document segmentation method according to claim 30, wherein said table type judging step comprises a syntax judging step, of judging the table type with reference to the cell position data, the cell vectors and the cell data generated in said table analyzing step, a partial character line extracting step, of extracting partial character lines from the cell data positioned at a particular position with reference to the cell position data and the cell data generated in said table analyzing step, and a similarity judging step, of judging the table type on the basis of similarity between the extracted partial character lines.

37.   (Previously Presented) A document segmentation method according to claim 23, further comprising a table reforming step, of reforming the table so that the number of cells in each column and each row becomes the same, by analyzing the table to be processed,

    wherein said table analyzing step includes analyzing the reformed table.

38. (Previously Presented) A document segmentation method according to claim 37, wherein said table reforming step comprises a supplementary data removing step, of removing data added to the table from the table data.

39. (Previously Presented) A document segmentation method according to claim 37, wherein said table reforming step comprises a multi-row/multi-column processing step, of reforming the table regularly by analyzing the structure of the table data.

40. (Previously Presented) A document segmentation method according to claim 37, wherein said table reforming step comprises a composite table processing step, of reforming the table by analyzing regularity of information description constituting the table.

41. (Previously Presented) A document segmentation method according to claim 37, wherein said table reforming step comprises:

a supplementary data removing step, of removing data added to the table from the table data; and

a multi-row/multi-column processing step, of reforming the table regularly by analyzing the structure of the table data.

42. (Previously Presented) A document segmentation method according to claim 37, wherein said table reforming step comprises:

- 14 -

a supplementary data removing step, of removing data added to the table from the table data; and

a composite table processing step, of reforming the table by analyzing regularity of information description constituting the table.

43.   (Previously Presented) A document segmentation method according to claim 37, wherein said table reforming step comprises:

a multi-row/multi-column processing step, of reforming the table regularly by analyzing the structure of the table data; and

a composite table processing step, of reforming the table by analyzing regularity of information description constituting the table.

44.   (Previously Presented) A document segmentation method according to claim 37, wherein said table reforming step comprises:

a supplementary data removing step, of removing data added to the table from the table data;

a multi-row/multi-column processing step, of reforming the table regularly by analyzing the structure of the table data; and

a composite table processing step, of reforming the table by analyzing regularity of information description constituting the table.

- 15 -

45.   (Currently Amended) A computer-readable storage medium storing a

document segmentation program for controlling a computer to perform document

segmentation for segmenting a list type table for describing data or a layout type table

representing between tags defined in a language for use in composing web pages, said program

comprising codes for causing the computer to perform:

a table analyzing step, of generating cell vectors representing characteristics of

cells having at least one of row width and column width of the cells, and cell position data

indicating a positional relationship between cells and cell vectors representing characteristics

of the cells, by analyzing a table pinched between a start tag and an end tag in a document to

be processed;

a table type judging step, of judging a table type of whether the table analyzed

in said table analyzing step is a list type table for describing data of a layout type table for

describing a layout of page with reference to the cell position data and the cell vectors

generated by said table analyzing step;

a first segment generating step, of generating a plurality of segments each of

which is pinched between the start tag and the end tag by dividing the table with a first method

in a case in which it is judged that the table [[type]] analyzed in said table analyzing step is a

list format type table in said table type judging step; and

a second segment generating step, of generating a plurality of segments each

of which is pinched between the start tag and the end tag by dividing the table with a second

method in a case in which it is judged that the table [[type]] analyzed in said table analyzing

step is a layout type table in said table type judging step.

- 16 -